

Using Regression Discontinuity to Estimate the Effects of a Tier 1 Research-Based Mathematics Program in Seventh-Grade

This is a pre-copyedited, author-produced PDF of an article that has published in *Exceptional Children* following peer review. The version of record [citation information below] is available online at <https://doi.org/10.1177/0014402918784541>

Jitendra, A. K., Harwell, M. R., Im, S-H., Karl., S. R., & Slater, S. C. (2018). Using regression discontinuity to estimate the effects of a Tier 1 research-based mathematics program in seventh-grade. *Exceptional Children*, 85(1), 46-65. First published August 3, 2018
<https://doi.org/10.1177/0014402918784541>

Asha K. Jitendra, Ph.D.
University of Minnesota

Michael R. Harwell, Ph.D.
University of Minnesota

Soo-hyun Im
University of Minnesota

Stacy R. Karl, B.S.
University of Minnesota

Susan C. Slater
University of Minnesota

Author's Note

Corresponding Author: Asha K. Jitendra, University of California, Riverside, Graduate School of Education, 900 University Ave., Sproul Hall 2127, Riverside, CA 92521. E-mail: asha.jitendra@ucr.edu

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant #R305A110358. The opinions expressed are those of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education. We gratefully acknowledge the efforts of all teachers and students who participated in the study.

Abstract

The present study used a regression discontinuity approach to test whether schema-based instruction (SBI) was effective for students identified with varying levels of mathematics difficulties (MD) who received Tier 1 instruction. The performance of these students was compared to similar students who received business as usual Tier 1 instruction. Results indicated SBI on average raised scores regardless of whether a student was categorized as having MD on outcome measures of mathematics problem solving, suggesting that SBI is effective for a wide range of proportional reasoning skills and general mathematical proficiency, which broadens the population of students that could benefit from SBI. Standardized slopes for the treatment effect also provide evidence of the effectiveness of SBI. Implications for practice and research are discussed.

Keywords. Regression discontinuity, proportional reasoning, mathematical problem solving, mathematics difficulties, schema-based instruction

Using Regression Discontinuity to Estimate the Effects of a Tier 1 Research-Based Mathematics Program in Seventh-Grade

Although the percentage of students reaching proficient levels in mathematics has increased over the past decade, it is disconcerting that the goal of the No Child Left Behind legislation to eradicate achievement differences among student subgroups has not yet been met (Dossey, Halvorsen, & McCrone, 2016). For example, the 2015 National Assessment of Educational Progress mathematics scores indicated that about two thirds (68%) of eighth-grade students with disabilities performed below the *Basic* level compared with 23% of students without disabilities (National Center for Education Statistics, n.d.). Thus, an important focus of research is to identify mathematics programs that are effective for students with mathematics difficulties (MD) as well as for their peers without MD, because students with MD often receive instruction in traditional mathematics classrooms.

One approach that is designed to promote positive mathematics achievement outcomes for all students uses a multi-tiered system of support (MTSS) model, a framework for providing high-quality instruction to all students and intervention support for some students (Fuchs & Fuchs, 2006; Fuchs & Vaughn, 2012; Vaughn & Swanson, 2015). The three-tier model is a common approach used to provide instructional support and services, with Tier 1 considered the key component of tiered instruction, wherein all students receive instruction within an evidence-based, scientifically researched core program that is typically aligned with state and national standards (e.g., Common Core State Standards [CCSS], National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Tier 2 includes targeted supplemental instruction provided in small groups to students who do not make adequate progress in Tier 1 and are considered at risk for academic failure. Tier 3 consists of intensive

individualized instruction for students who demonstrate inadequate response to Tier 2 interventions and are considered to be high at risk for failure.

The focus of this study is on Tier 1 instruction to improve seventh-grade students' proportional reasoning, a topic of particular importance in middle grades. Proportional reasoning requires understanding the concept of ratios and that two or more ratios are equal. It also requires the ability to extract relevant information to develop a representation of the problem situation, which is challenging for many children and adolescents (Özgün-Koca & Altay, 2009). It is well documented that many students' difficulties with proportion and proportion-related tasks may be due to their lack of understanding of ratios. A growing body of evidence suggests that proportional reasoning and understanding fractions predict later mathematics achievement (Bailey, Hoard, Nugent, & Geary, 2012; Siegler et al., 2012; Siegler, Fazio, Bailey, & Zhou, 2013). In sum, if we can achieve high rates of student success for *all* students with respect to developing proportional reasoning through Tier 1 supports such as differentiated instruction then there could be savings and redeployment of instructional resources typically used in Tier 2 because fewer students will need more intensive intervention.

The purpose of the present study was to use a regression discontinuity (RD) approach to examine whether a research-based instructional program, schema-based instruction (SBI), was effective for students with and without MD who received Tier 1 instruction in proportional reasoning. The performance of these students was compared to similar students who received business as usual instruction. RD designs are quasi-experimental and permit strong causal inferences like those associated with randomized controlled trials for research in tiered instruction (Ashworth & Pullen, 2015; Lee & Lemieux, 2010). In special education, RD designs are appropriate in studying the effectiveness of interventions such as with MTSS studies when

using a cut-off so that “the treatment effect observed visually around a cut-off value can also be extended in both directions from the cut-off value” (Ryoo & Pullen, 2017, p. 138).

Research on Tier I Instruction in Proportional Reasoning

A review of research on proportional reasoning for middle school students that included comprehensive coverage of topics (e.g., ratios and proportions, scale drawings, percent and percent of change) found only a few randomized controlled Tier 1 studies (Jitendra et al., 2015; Jitendra, Harwell, Im, Karl, & Slater, 2018; Jitendra, Harwell, Karl, Simonson, & Slater, 2017; Jitendra, Star, Dupuis, & Rodriguez, 2013; Jitendra, Star, Rodriguez, Lindell, & Someki, 2011). These studies were conducted in one upper Midwest state (Jitendra et al., 2011, 2013, 2015) and two states in the southeastern and western regions of the U.S. (Jitendra et al., 2018; Jitendra, Harwell, Karl, et al., 2017) and tested the efficacy of SBI, an instructional program designed to help seventh-grade students make sense of their reasoning related to proportions in word problem contexts. SBI is an instructional approach that has its roots in schema theory and research on expert problem solvers and is guided by cognitive models of mathematical problem solving (see Marshall, 1995; Mayer, 1999). Essential features of SBI include identifying the underlying problem structure, using visual-schematic representations that illustrate the mathematical relations among key elements in the problem, facilitating problem solving, including developing procedural flexibility, and developing metacognitive strategy skills (see Woodward et al., 2012). In addition, SBI integrates effective instructional practices (e.g., explicit modeling, scaffolding instruction with sufficient examples) that improve problem-solving performance of students with MD (see Gersten et al., 2009).

In these studies, teachers randomly assigned to SBI classrooms received professional development to implement SBI five days a week for approximately 45-50 min over a 6-week

period to teach proportional reasoning skills, whereas teachers randomly assigned to the control condition taught the same topics from district-adopted mathematics textbooks. In the initial efficacy trial of SBI, Jitendra et al. (2011) randomly assigned 21 seventh-grade classrooms sampled from three schools across two districts to one of three conditions: (1) SBI only, (2) SBI + Tutoring, or (3) business-as-usual control. However, the wide variation in implementation of tutoring across the two districts was problematic in terms of inclusion of SBI + Tutoring as a distinct condition; thus, the two SBI conditions were combined for data analysis. Multilevel statistical analyses results showed the posttest difference on the proportional problem-solving test (PPS) favoring the SBI group was statistically significant (standardized effect size $g = 0.32$), but the effects of SBI were not maintained at a 4-week follow-up (delayed posttest, $g = 0.22$, ns). The latter finding was likely due to a lack of power to detect significant differences given the modest number of classrooms ($j = 21$). This study did not report results separately for students with and without MD.

Subsequent studies of SBI improved the research design by increasing the sample size to include more classrooms ($j = 42$ in Jitendra et al., 2013; $j = 82$ in Jitendra et al., 2015; $j = 59$ in Jitendra et al., 2018) across more schools ($k = 6$ in Jitendra et al., 2013; $k = 58$ in Jitendra et al., 2015; $k = 36$ in Jitendra et al., 2018), reducing the direct involvement of the research staff in SBI implementation (Jitendra et al., 2013, 2015, 2018; Jitendra, Harwell, Karl, et al., 2017), and randomly assigning teachers to SBI or control and then randomly selecting one of their classrooms to participate in the study (i.e., Jitendra et al., 2015, 2018; Jitendra, Harwell, Karl, et al., 2017), meaning that each teacher taught in a SBI or control classroom but not both. The two recent studies of SBI (Jitendra et al., 2018; Jitendra, Harwell, Karl, et al., 2017) used similar conditions, materials, and methods in different geographic locations and in schools more

demographically diverse than in Jitendra et al. (2015), which involved a rigorous replication of earlier studies of SBI and included a sample of students ($N = 1,981$) larger than samples ($N = 436$ and 1,163) in Jitendra et al. (2011, 2013).

Most studies of SBI reported results separately for students with and without MD (see Jitendra, Dupuis, Star, & Rodriguez, 2016; Jitendra, Harwell, Dupuis, & Karl, 2017; Jitendra, Harwell, Karl, et al., 2017). Taken together, findings suggested that SBI improved the proportional problem-solving performance of all students, including students with MD, on both PPS immediate (effect size range: 0.36 to 0.63) and delayed posttests (effect size range: 0.29 to 0.33). With regard to solving transfer problems (e.g., probability), there were no SBI effects in Jitendra et al. (2013; see also Jitendra et al., 2016) on a researcher-developed transfer measure that included items not directly aligned with the taught content. Jitendra et al. (2015) reported no SBI effects on the Process and Applications subtest of the Group Mathematics Assessment and Diagnostic Evaluation [GMADE] (Pearson Education, 2004), a standardized test that assessed overall mathematical problem solving involving multiple content areas (e.g., algebra, geometry). In Jitendra, Harwell, Karl, et al. (2017), posttest differences favoring the treatment group were statistically significant for the GMADE posttest for the full sample (standardized effect size $g = 0.32$), but not for students with MD. This study resembled a replication but sample size limitations ($j = 20$ teachers/classrooms and their 373 students from 10 middle schools) suggest it is more appropriately characterized as a modest follow-up to Jitendra et al. (2015) with limited generalizability.

In the recent replication study (Jitendra et al., 2018), which is an extension of the rigorous large-scale, randomized controlled efficacy study conducted by Jitendra et al. (2015), the authors examined whether the impact of SBI would replicate in different geographical regions of the

country and in schools more demographically diverse than in the Jitendra et al. (2015). Results revealed statistically significant differences between conditions on proximal (effect sizes of 0.47 and 0.29 on the PPS posttest and delayed posttest) and distal measures (effect size of 0.31 on the GMADE) of mathematics problem solving, with effects sizes for proximal measures similar to those reported in Jitendra et al. (2015). However, this study did not test whether SBI enhances proportional reasoning skills for a wide range of mathematical proficiencies (i.e., students with and without MD).

Study Context and Rationale for a Regression Discontinuity Design

Jitendra, Harwell, Dupuis, et al. (2017) examined the effectiveness of SBI for a subsample of students with MD selected from Jitendra et al. (2015). The GMADE (Pearson Education, 2004) score in the student sample corresponding to the 35th percentile on a general measure of mathematical problem solving (Process and Applications subtest) was used to categorize students as having MD. GMADE pretest scores at or below 12 led to a categorization as having MD and scores larger than 12 to a categorization as not having these difficulties. Jitendra, Harwell, Dupuis, et al. (2017) fitted two-level students-within-classrooms multilevel models to posttest and delayed posttest data capturing proportional reasoning skills (PPS) for a sample of MD students in separate analyses. The multilevel models used the student covariates PPS pretest, race, and gender and classroom covariates teacher gender and experience, percentage of students receiving special education services, percentage of students eligible for a free/reduced price lunch (FRL), and a treatment variable (1 = SBI, 0 = Control). Based on a sample of 806 students classified as MD clustered within 82 classrooms, Jitendra, Harwell, Dupuis, et al. (2017) reported SBI improved student scores on the PPS posttest and delayed

posttest compared to a control condition, suggesting SBI could be used effectively for students classified as having MD in the short and long term.

The results of Jitendra, Harwell, Dupuis, et al. (2017) raise an important question: What is the range of proportional reasoning skills and general mathematical proficiencies for which SBI enhances student performance? For example, evidence that SBI enhances proportional reasoning skills for a wide range of mathematical proficiencies (e.g., students with and without MD, students well below the MD cutoff) further broadens the population of students SBI could be used with; correspondingly, evidence that SBI is primarily effective for students with a limited range of mathematical proficiency (e.g., students just below the MD cutoff) narrows the potential population of students that could benefit from SBI. Taking student background variables and classroom factors into account when examining the effectiveness of SBI for a range of proportional reasoning skills and general mathematical proficiencies also enhances the utility of SBI.

The multilevel model of Jitendra, Harwell, Dupuis, et al. (2017) could be used to predict outcome values reflecting the impact of SBI for a range of proportional reasoning skills and general mathematical proficiencies that take student background variables and classroom factors into account by examining different combinations of covariate values. For example, predicted PPS posttest scores could be generated for PPS pretest (score range 0-31, total of 32 possible values) by adding this student covariate to a model containing the SBI predictor (1 = SBI, 0 = control). This would require fitting 64 statistical models (32 for students in the SBI condition and 32 for those in the control condition assuming sufficient data for these analyses were available) and produce 64 sets of predicted PPS posttest scores (e.g., students with a pretest score of 0 versus 1 versus 2, etc.). Variation among the sets of predicted PPS posttest scores for the SBI and

control conditions speaks to questions about the range of proportional reasoning skills and general mathematical proficiencies SBI is effective for. To learn whether patterns in the 64 sets of predicted PPS posttest scores are consistent across student background variables and classroom factors requires fitting additional models.

For example, if teacher gender is added to a model containing student PPS pretest scores a total of $2 \times 64 = 128$ models would be fitted. Adding additional covariates such as student race, years of teaching experience, and percentage of students receiving special education services in a classroom creates an unmanageable number of models even if the number of values in the predictive models is limited on some basis (e.g., collapsing years of teaching experience into quartiles such as 1-5 years, 6-10, 11-15, >15). This prompted us to seek an alternative method for exploring the range of proportional reasoning skills and general mathematical proficiencies for which SBI enhances student performance and its consistency across student background variables and classroom factors.

We turned to a RD approach because it provides direct information about the effectiveness of SBI without requiring analyses of hundreds of statistical models and because of its ability to support strong causal inferences (What Works Clearinghouse, 2014). Our use of RD compares SBI and control conditions on the PPS posttest, PPS delayed posttest, and the GMADE posttest for varying pre-intervention levels of proportional reasoning and general mathematical proficiencies. Like Jitendra, Harwell, Dupuis, et al. (2017), we used the GMADE pretest score corresponding to the 35th percentile as an MD cutoff and the GMADE posttest as the outcome; unlike Jitendra, Harwell, Dupuis, et al. (2017), we also used the PPS pretest score corresponding to the 35th percentile as an MD cutoff for the PPS posttest and delayed posttest outcomes and included students with and without MD in the analyses.

Although previous studies have evaluated the impact of Tier 2 interventions using a RD design (e.g., Baker, Smolkowski, Chaparro, Smith, & Fien, 2015; Bryant et al., 2008; Bryant, D. P., Bryant, B. R., Gersten, Scammacca, & Chavez, 2008), we found no previous studies using this design to evaluate the impact of a Tier 1 instruction to improve the mathematical performance for students with a wide range of mathematical proficiencies. RD designs have commonly been used to evaluate the effects of Tier 2 interventions when all students with special needs receive treatments, and the RD approach is used to detect effects of intervention even without the use of a control group (Ashworth & Pullen, 2015). In addition, studies have used student scores on specific measures as the cutoff to determine which students were at risk for academic failure and needed an additional tier of instruction or whether a more intensive secondary tier (received Tier 1 and Tier 2) would improve academic outcomes more so for students who received only Tier 1 (Baker et al., 2015). Our RD approach differs from the RD designs traditionally used in special education where “the focus is on the discontinuity of a dependent variable around the specific value in an independent variable” (Ryoo & Pullen, 2017, p. 138). We build on the traditional RD approach by using local linear regression models to estimate the treatment effect for varying bandwidths (collections of pretest scores) “to establish robustness of the effect size estimate” (Ryoo & Pullen, 2017, p. 141).

The logic of RD is typically based on the crucial role of pretests in taking student differences into account (Steiner, Cook, Shadish, & Clark, 2010), in that students with similar pretest scores can often be treated as approximately equal on background variables such as those capturing student gender and race, enhancing causal inferences (Bloom, 2010). We followed the example of Robinson (2010) in employing RD and note that dependencies among outcomes for students sharing a teacher are not taken into account in this approach, which risks biasing the

findings to some extent. Comparing the RD treatment effects with those reported in Jitendra et al. (2018) whose study of SBI took dependencies into account will provide insight into its likely impact; a pattern of effect sizes that are consistent for the RD analyses and Jitendra et al. (2018) provides evidence dependencies have not seriously biased estimates of treatment effect whereas a conflicting pattern suggests dependencies may be seriously biasing the RD results.

Method

Participants and Setting

The sample of students ($n = 1,078$ for the PPS pretest; $n = 1,120$ for the GMADE pretest) was taken from the Jitendra et al. (2018) replication study, which randomly assigned classrooms to SBI or control conditions and assumed a classroom's treatment status was unrelated to whether students within a classroom were subsequently categorized as having MD. For this sample, a score of 9 or less on the PPS pretest led to a student being categorized as having MD and scores greater than 9 as not having MD; for the GMADE posttest a score of 10 or less on the GMADE pretest led to a student being categorized as having MD and scores greater than 10 as not having MDs. Table 1 summarizes demographic information about teachers and students for the current study.

Students. The majority of students were White (53.5%), with 27.9% Hispanic, 9.1% Black, 4.9% Asian, and 3.3% Multiracial. The mean age of these students was 12 years, 7 months ($SD = 5$ months). Approximately 9.8% percent of the student sample received special educational services and 10.3% were English language learners.

Teachers. The 59 participating seventh-grade mathematics teachers' (49 females) experience ranged from 1 to 41 years ($M = 10.4$ years; $SD = 8.6$). All teachers were certified to teach mathematics, 12 were certified in all subjects (generalist), 2 were also certified to teach

science, and 15 were certified in subjects other than mathematics or science. Similarly, all teachers were certified to teach grades 6-8; 34 teachers were also certified to teach grades 9-12 and 18 were certified to teach grades K-5. The majority of participating teachers were White (89.8%), with three Hispanic (5.1%), two Black (3.4%), and one Asian teacher (1.7%). Fifty-two (88.1%) teachers taught in suburban schools, 6.8% in rural schools, and 5.1% in urban schools.

Research Design

Jitendra et al. (2018) used a prospective randomized cluster design with both longitudinal (pretest, posttest, delayed posttest) and cross-sectional data in which teachers/classrooms served as clusters. One class of students for each of 59 teachers sampled from five districts in two states in the U.S was randomly selected to participate: 27 teachers were from urban/suburban/rural districts in a southeastern state and 32 from a western state. Forty-one teachers and their class were randomly assigned to treatment (SBI) or control conditions regardless of their district or geographic location.

The remaining 18 teachers were in eight schools in a district that mandated participation in the study. Teachers in this district were organized into grade and content level groups to work on the district's Professional Learning Community (PLC) teams. As such, random assignment was done at the school level to minimize impact on the district's PLC model, which required treatment and control teachers in a school to share materials and instructional practices, raising concerns effects could be contaminated. With school-based random assignment of 18 teachers to the SBI and control conditions, there were a total of 34 treatment classrooms and 25 control classrooms. Jitendra et al. (2018) assessed dependency among classrooms within schools in the data analyses by adding a predictor indicating the nature of the random assignment (classroom vs. school-based) and found this predictor was not statistically significant in any analysis.

Description of the Treatment Instruction

Teachers in the SBI condition replaced the lessons on ratio/proportion and percent in their district curriculum with the SBI program. The SBI program includes 21 lessons that can be completed in about 30 days (some lessons take more than a day to implement). Organized into two units of 10 lessons each, the first unit, Ratio/Proportion, focuses on the meaning of ratios, equivalent ratios, and rates, as well as word problems involving ratios, proportions, and scale drawings. The second unit, Percent, focuses on the meaning of percent (as well as fractions and decimals as alternative representations) and word problems involving part-whole comparisons, percent of change problems, including those involving sales taxes, discounts, tips, and simple interest, as well multistep adjustment percent of change problems. The last lesson in each unit presents real-world scenarios (i.e., designing a recording studio, constructing a digital planetarium) with a focus on students working in small groups to solve a variety of problems involving ratios and proportional relationships. Lesson 21 in the SBI program provides a review of the content from both units.

The 21 lessons are highly specified such that a detailed teacher guide (see Appendix A, sample excerpts of scripts for solving ratio and proportion problems in Lessons 3 and 6 located in the online supplemental materials) supports teachers in implementing proportional reasoning activities (see “Professional Development”). Instructional practices include whole class instruction followed by partner or small group work using a Think-Plan-Share strategy to enhance students’ critical thinking. The SBI program incorporates effective instructional practices (e.g., teacher modeling problem solving procedures by thinking aloud, using prompts to help clarify and refine student thinking) and includes differentiated instruction to meet the diverse needs of students. For example, challenge problems (see Appendix A, sample challenge

problems from the SBI program located in the online supplemental materials) that involve procedures with connections tasks are included that emphasize the use of procedures but require engaging in a productive struggle to build connections to prior relevant knowledge (Smith & Stein, 2011). These problems, however, are assigned only to students who have the fundamental conceptual knowledge. In addition, homework complements and reinforces critical concepts and skills taught in the SBI program. The cognitive demands of homework problems also vary and are assigned based on student needs.

The key features of the SBI program are described in Jitendra et al. (2015) and are aligned with the recommendations articulated in the What Works Clearinghouse's research synthesis on improving students' mathematical problem-solving performance (Woodward et al., 2012). The problem-solving knowledge application activity in the SBI program is designed to apply and extend understanding of critical content (e.g., ratios/rates, percent) to solve problems involving ratios and proportional relationships. Students learned to apply problem-solving procedures for a given class of problems (e.g., ratio, proportion, percent of change), including checks to monitor and reflect on the problem-solving process (e.g., when, how, and why to use multiple strategies—equivalent fractions, unit rate, cross multiplication). Further, students engaged in discourse around the content as teachers guided them to: (a) recognize the problem type (i.e., ratio, proportion, or percent), (b) connect the problem to previously solved problems by thinking about how problems within and across types are similar or different, (c) identify and represent critical information in the problem using a visual-schematic diagram that illustrates the relationships between relevant quantities in the problem, (d) estimate the answer, (e) select an appropriate strategy (i.e., unit rate, equivalent fractions, or cross products) based on the

quantities in the problem, (f) solve and present the solution within the context of the problem, and (g) check the reasonableness of the solution.

Professional development (PD). The 2-day PD on proportional reasoning took place before the start of the study at each site (between August and December). The PD leader, an experienced PD trainer, curriculum developer, and an expert in problem solving who was external to the project conducted the training at each district site. The PD training focused on developing awareness of the importance and complexity of proportional reasoning and of the ways students understand and learn, as well as discussing the background and principles of the of the SBI and orienting teachers to SBI program design in terms of lesson structure and scope and sequence. In addition, the PD covered critical SBI practices (e.g., recognizing problem types, generating estimates, applying multiple-solution strategies) to support student learning of ratio, proportion and percent and provided teachers with guidance for implementing SBI instructional practices by having them view multiple short video clips of lesson elements implemented by teachers in the initial study. The PD facilitator discussed not only ways to set up and manage materials, but also pacing of the lessons to meet the diverse needs of students in their classrooms.

“Business as usual” (BAU) control instruction

Teachers assigned to the BAU condition provided instruction that would typically occur in a seventh-grade math class, addressing the same topics (i.e., ratio, proportion, and percent) and content (aligned with state standards) taught in the treatment classes for the same period of time (45 minutes, on average, daily for 6 weeks).

We gathered information on textbooks used in the control classrooms from a written teacher questionnaire, which showed that teachers in the control classrooms used four different textbooks from the following publishers: Big Ideas Learning; Houghton Mifflin Harcourt;

McGraw-Hill; and Pearson Education. In general, the four textbooks addressed practices outlined in the CCSS standards and instruction in these programs ranged from a balanced instructional approach (discovery learning and scaffolded instruction) to digital interactive learning. A review of instructional features in the control classroom textbooks suggested that they do not consistently and extensively overlap with those in the SBI program in ways that would distort estimates of the effects of SBI (see Table 2).

Measures

Students in the treatment and control conditions completed the PPS test and GMADE prior to and immediately following the treatment. In addition, the same PPS test used for pretest and posttest was given 9 weeks after the end of the treatment (delayed posttest). Classroom teachers administered these assessments following standardized protocols. Although the tests were untimed, each test could be completed in 50 minutes on average. We also collected data for student race, sex, special education, ELL, and FRL status.

PPS test. The PPS test, which included 22 multiple-choice questions and three short-response problems with four open-ended items, measured students' ability to solve proportion problems involving ratios/rates and percent (see Appendix B, sample items located in the online supplemental materials). This test was developed using released items from NAEP and TIMSS as well as questions from past state mathematics assessments. Each multiple-choice item was scored as correct or incorrect. Each open response item was scored using a rubric, which emphasized correct reasoning; these responses were scored on a 0-to-2 point scale. We calculated students' score on the PPS test by computing the sum of the total points earned; the short-response and multiple-choice items contributed unequally to the overall score. The average inter-rater reliability for the short answer items using an intra-class correlation produced values

of 0.88, 0.91, and 0.92 at pretest, posttest, and delayed posttest, respectively. Internal consistency estimates of the PPS using the jMetrik software (Version 2.1.0; Meyer, 2011) produced coefficient omega (Dunn, Baguley, & Brunsden, 2013) values of 0.77, 0.84, and 0.84 for the PPS pretest, posttest, and delayed posttest, respectively.

GMADE. The 30-item Process and Applications subtest of the GMADE (Pearson, 2004), a norm-referenced standardized assessment was used to sample students' mathematical competency, specifically, their understanding of the process for solving problems as well as mathematical language and concepts and their ability to apply relevant operations to solve word problems across multiple content areas (e.g., algebra, geometry, number and operations). All items were scored as correct or incorrect. The coefficient omega estimates for this sample were 0.69 for the pretest and 0.76 for the posttest.

Fidelity of implementation

Two classes per treatment teacher and one class per control teacher were videotaped to collect data on proportional problem-solving instruction. Trained coders (graduate assistants, project coordinator) assessed procedural fidelity and adherence to SBI by observing videotaped lessons using the same rubric as Jitendra et al. (2015) to document the presence or absence of key instructional features of SBI (e.g., identify the problem type, represent critical information in the problem text using an appropriate diagram, discuss multiple solution strategies, check the solution). This rubric with items evaluated on a 0-to-3 scale (0 = did not implement to 3 = high level of implementation) was also used in the control condition to assess program differentiation and determine whether control teachers spontaneously provided instruction that was similar to the key elements of SBI (Dane & Schneider, 1998). In addition, the quality of instructional delivery of lessons in both treatment and control classes were evaluated on the same 0-to-3 scale

and assessed teachers' ability to clarify the lesson purpose, provide lesson closure, manage instructional time, and minimize mathematical errors.

Two of four coders independently assessed fidelity for each classroom producing 186 ratings (68 ratings in treatment classrooms; 25 in control classrooms). Estimated inter-rater reliability via intra-class correlations for the ratings averaged 0.92 (range 0.84 to 0.98) for procedural fidelity and 0.91 for quality (range 0.87 to 0.94).

Analysis of fidelity. Table 3 shows fidelity of implementation and quality of instruction data. Results of *t*-tests indicated statistically significant and substantial differences between the treatment and control groups on the total fidelity of implementation score, $t(57) = 8.29, p < .001$, with treatment teachers ($M = 15.71; SD = 2.42$) implementing SBI elements more than control teachers ($M = 10.16; SD = 2.70$). These data provide evidence of program differentiation (Dane & Schneider, 1998) in that there were clear differences in SBI instructional elements across the two groups. As expected, there were no statistically significant differences between groups, $t(57) = 0.45, p = 0.69$, on the quality of instruction (SBI: $M = 8.75, SD = 1.06$; Control: $M = 8.56, SD = 2.16$).

Results

The analyses compared SBI and control students using normal-theory-based multiple regression in which a selected range of pretest scores defined the sample and treatment (SBI = 1, Control = 0) served as a covariate, with PPS posttest, PPS delayed posttest, and the GMADe posttest serving as outcomes. Each outcome was analyzed separately using the SPSS 22.0 software package (IBM Corp., 2015).

We initially performed descriptive analysis that included examining the correlations among all measures. Table 4 shows correlations among variables by mathematical difficulties

status and treatment group. The descriptive statistics in Table 1 summarize student characteristics for our sample whose GMADE pretest scores ranged from 5-15 (bandwidth = ± 5) and PPS pretest scores ranged from 4-14 (bandwidth = ± 5) (see Table 6). Following the example of Robinson (2010) we initially conducted two-sample *t*-tests to learn whether SBI and control conditions produced similar outcomes for students with PPS and GMADE pretest scores right at the cutoff of MD status, i.e., 9 (MD) or 10 (almost MD) for the PPS pretest and 10 (MD) or 11 (almost MD) for the GMADE pretest. The use of a RD design strengthens inferences because students with similar PPS pretest scores (9 or 10) or GMADE (10 or 11) should be quite similar and minimize the impact of omitted variables (Bloom, 2010).

Table 5 shows statistically significant differences between SBI and control conditions for students whose PPS pretest score was 9 or 10 on the PPS posttest ($t(252) = 5.068, p < .001, d = 0.64SD$) and delayed posttest ($t(243) = 3.340, p = .001, d = 0.43SD$) (samples are small because only students with specified pretest scores were included), with SBI students outperforming control students. For the GMADE posttest there was a statistically significant difference between SBI and control conditions for students whose GMADE pretest score was 10 or 11 ($t(235) = 2.746, p = .006, d = 0.36SD$), with SBI students outperforming control students. These findings suggest that the impact of the treatment on all three outcomes was similar for students categorized as having (or almost having) MD.

The RD results for the PPS posttest and delayed posttest and the GMADE posttest are reported in Table 6 and Figure 1. These analyses are based on different bandwidths of the PPS and GMADE pretests. There was a significant effect of SBI on PPS posttest and delayed posttest scores for every bandwidth of the PPS and GMADE pretests studied. For example, the multiple regression for a bandwidth of ± 1 only used students whose PPS pretest score was 8, 9, or 10 ($n =$

376) and found a significant treatment effect (slope) of 2.11 which is the estimated discontinuity (vertical difference). That is, for students with PPS pretest scores of 8, 9, or 10, SBI students on average scored 2.11 points higher on the PPS posttest than control students ($p < .001$, standardized slope = $0.225 SD$ which estimates the difference between SBI and control with the selected band of pretest scores (8, 9, 10) taken into account). Note that this finding ideally approximates those obtained in a randomized design, meaning student background variables and classroom factors are not expected to bias the treatment effect.

The 2.11 difference is represented in Figure 1 by the discontinuity between the two bolded lines for the PPS pretest score of 9; the gray shading represents the 95% confidence interval of best fit. The bolded lines in Figure 1 represent the estimated PPS posttest score at the median of the selected bandwidth of PPS pretest scores, and as sample size shrinks the accuracy with which posttest scores are estimated using the fitted regression model declines for more extreme PPS pretest values and thus the shaded area widens. The fact there were significant treatment effects for every bandwidth studied for the PPS posttest and delayed posttest suggests the SBI program is effective for a relatively broad range of proportional reasoning skills for this outcome. Treatment effects also appeared for all bandwidths studied for the GMADE posttest, suggesting the SBI program on average again was effective for a relatively broad range of general mathematical proficiencies.

The standardized effect sizes in Table 6 are smaller than the Cohen's d values reported in Jitendra et al. (2018) for the PPS and GMADE outcomes, which is likely due to a narrower range of pretest scores used in each RD analysis. Importantly the RD results in Table 6 are consistent with Jitendra et al. (2018) in that SBI-control differences are all statistically significant and always favor SBI. This pattern suggests dependencies in the outcomes did not seriously bias the

RD results. The RD results also provide evidence that SBI enhances proportional reasoning skills for a wide range of mathematical proficiencies (i.e., students with and without MD), which broadens the population of students for whom SBI would be beneficial.

Discussion

In the previous SBI study (Jitendra, Harwell, Dupuis, et al., 2017), students with MD who received Tier 1 instruction outperformed students with MD in a business-as-usual condition. However, the sampled schools, teachers, and students were from one upper Midwest state, which included mostly rural schools and a sample whose distributions of key student background variables (e.g., percentage of White students, ELL students, and students eligible for FRL) were below national averages (U.S. Department of Education, 2015). Additionally, the results of Jitendra, Harwell, Dupuis, et al. (2017) raised the question of how effective SBI is for students with a range of proportional reasoning skills and a range of general mathematical proficiencies (GMADE). Thus, the purpose of this study was to test the impact of a Tier 1 instructional program (SBI) for students identified with varying levels of MD in geographic locations outside the upper Midwest with a sample whose distributions of key student background variables are closer to U.S. averages. Using a regression discontinuity approach, this study provided evidence SBI is effective for a wide range of proportional reasoning skills and general mathematical proficiency and thus broadens the population of students SBI could be used effectively with. Standardized slopes for the treatment effect also provide evidence that the magnitude of the SBI effect (about half a standard deviation) is substantial.

Evidence of SBI Tier 1 Impact for Students with MD

The major finding was that the SBI program, which targeted understanding of ratios and proportional relationships, significantly improved proportional reasoning outcomes for students

categorized as having MD (or almost having MD as well as students further from the cut score) at both immediate and delayed PPS posttests. SBI students with MD performed approximately two-thirds of a standard deviation higher than students receiving typical mathematics instruction in the same content on the PPS at immediate posttest. This translates to approximately 74% of treatment students scoring above the mean of control classrooms (Lipsey et al., 2012). A second finding was that SBI students with MD maintained these gains nine weeks following the treatment. SBI students with MD performed almost half a standard deviation higher than students receiving typical mathematics instruction on the PPS at delayed posttest, meaning that approximately 67% of treatment classrooms scored above the mean of control classrooms (Lipsey et al., 2012). Furthermore, the findings of the RD analyses are noteworthy given the significant treatment effects for SBI students with a wide range of incoming (pretest) proportional reasoning skills on the PPS posttest and delayed posttest. That is, the positive impact of SBI on improving proportional reasoning skills and general mathematical proficiencies was insensitive to student's initial proportional reasoning skills and general mathematical proficiencies.

Overall, our findings provide strong evidence that the SBI program's focus on research-based instructional practices (e.g., recognizing the problem structure and connecting the problem to previously solved problems, representing the problem situation using an appropriate visual-schematic diagram that shows the relationships between relevant quantities, monitoring and reflecting on the problem solving processes) was effective in enhancing the proportional reasoning performance of students with MD. These findings build on and expand previous results showing that the SBI program can effectively support students with MD in developing

proportional reasoning skills (see Jitendra et al., 2016; Jitendra, Harwell, Dupuis, et al., 2017; Jitendra, Harwell, Karl, et al., 2017).

In addition, the finding on the GMADE is encouraging regarding SBI students' improved performance in overall mathematical problem solving. SBI students with MD performed close to half a standard deviation higher than students receiving typical mathematics instruction on the GMADE posttest. This translates to approximately 64% of treatment students scoring above the mean of control classrooms (Lipsey et al., 2012). Additionally, the findings of the RD analyses are noteworthy given the significant treatment effects for every bandwidth studied for the GMADE posttest. These findings on the GMADE are not consistent with previous Tier 1 SBI studies (Jitendra et al., 2015; Jitendra Harwell, Karl, et al., 2017). It is worth noting that the PD training provided to teachers in Jitendra et al. (2015) differed from the Jitendra Harwell, Karl, et al. (2017) and current studies with regard to the sample size of teachers participating in the training. The small group of teachers at each site in Jitendra Harwell, Karl, et al. (2017) and current studies possibly led to richer interactions between the PD leader and participants, with the PD leader expanding on the problem solving focus in Jitendra et al. (2015) in a way that the leader was more responsive such as clarifying materials in context to improve teacher practices related to problem solving that, in turn, led to greater gains in general problem solving achievement for students in the current study. The SBI effect in Jitendra, Harwell, Karl, et al. (2017) was not statistically significant possibly due to the modest number of classrooms ($j = 20$); however, the effect size of $g = 0.26$ is non-negligible (see What Works Clearinghouse, 2014). Thus, one possible explanation for the positive SBI effect in the current study, with adequate statistical power, is that the SBI program's emphasis on priming the problem structure and connecting new learning (to be solved problems) to previously solved problems and the use of

multiple exemplars with attention on understanding of mathematical concepts (e.g., ratio, proportion, percent) enabled students with MD to adequately transfer the previously acquired knowledge. In short, teacher time dedicated to proportional problem solving possibly created bridges between this learning, which focused on the use of procedures that developed understanding of mathematical concepts and ideas, and general mathematical problem solving for students with MD.

Implications for Practice

Within the context of prevention science and MTSS, there is evidence to consider SBI to be efficacious and effective based on this study and multiple randomized controlled studies that produced statistically significant short and long-term effects for students with and without MD when SBI was implemented in real-world classrooms (Schulte, 2016). As such, SBI within a school could be used as Tier 1 instruction in MTSS to address the mathematics outcomes of students with and without MD. The study has important implications for policy makers and teachers in that SBI with its focus on effective instructional practices integrated with differentiated instruction in Tier 1 can be used to meet the needs of a range of learners, including students with MD with regard to solving problems involving ratios, proportions, and percent. Transfer to solving novel problems involving varying topics (e.g., algebra, geometry) was also feasible given that teachers in this study were provided guidance (i.e., scaffold student learning) in reaching a wide range of students to apply the SBI program components in multiple contexts with attention to understanding critical mathematical concepts. Specifically, teacher training focused on helping students to recognize the mathematical problem structure via schematic diagrams using problem solving and metacognitive strategy instruction and encouraging students to employ multiple solution methods is important in enhancing proportional reasoning and

overall mathematical problem solving. Furthermore, as increasing number of districts and schools throughout the country are using tiered instructional approaches, RD would be a useful approach to evaluate instructional programs within a MTSS in terms of whether students with varying levels of MD improve relative to comparable peers.

Limitations and Future Research

The study has some limitations that are important in considering the findings in this study. First, the SBI program included multiple components, and we do not know which of the components contributed to the outcomes and which may be less influential. As such, there is a need “to experimentally manipulate and isolate the impact of the various components determining their relative effects” (Vaughn et al., 2011, p. 959). However, based on program differentiation results from the fidelity data it seems that treatment teachers were more likely than control teachers to implement SBI lesson elements that are relatively unique to SBI (i.e., identify the problem type by focusing on the key problem features, connect the new problem to previously solved problems, generate an estimate prior to solving the problem, discuss multiple solution strategies, evaluate the solution) and that control teachers were most likely to implement instructional practices that most mathematics teachers typically engage in (e.g., solve the problem and present the solution within the context of the problem). These findings suggest that each of SBI elements implemented significantly more (i.e., moderate to high level of implementation) by treatment teachers may have collectively contributed to positive student outcomes.

A second limitation concerns the measures used to identify students as having MD. The PPS and GMADE tests are both psychometrically sound instruments, but choice of cut-score for identifying MD students requires additional research to providing validity evidence for these test

values. In addition, the GMADE is also lengthy and expensive for most schools to use as a screening measure.

Conclusion

The CCSS in mathematics require middle school teachers to focus sufficient instructional time on proportional reasoning to assure its careful development. The SBI program makes an important contribution to the field by indicating it is possible for teachers to accomplish this in inclusive classrooms composed of students with and without MD. Additionally, this and prior research (e.g., Jitendra et al., 2015, 2018) indicate that “professional development focused on both enactment of curricular replacement units and effective teaching” of students with MD and their classmates without MD may enhance student outcomes (August et al., 2014, p. 79).

References

Ashworth, K. E., & Pullen, P. C. (2015). Comparing regression discontinuity and multivariate analyses of variance: Examining the effects of a vocabulary intervention for students at risk for reading disability. *Learning Disability Quarterly*, 38, 131–144.
doi:10.1177/0731948714555020

August, D., Branum-Martin, L., Cárdenas-Hagan, E., Francis, D. J., Powell, J., Moore, S., & Haynes, E. F. (2014). Helping ELLs meet the common core state standards for literacy in science: The impact of an instructional intervention focused on academic language. *Journal of Research on Educational Effectiveness*, 7, 54–82.
doi:10.1080/19345747.2013.836763

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology*, 113, 447–455. doi:10.1016/j.jecp.2012.06.004

Baker, S. K., Smolkowski, K., Chaparro, E. A., Smith, J. M., & Fien, H. (2015). Using regression discontinuity to test the impact of a Tier 2 reading intervention in first grade. *Journal of Research on Educational Effectiveness*, 8, 218–244.
doi:10.1080/19345747.2014.909548

Bloom, H. (2010). *Modern regression discontinuity analysis*. New York, NY: MDRC.

Bryant, D. P., Bryant, B. R., Gersten, R., Scammacca, N., & Chavez, M. M. (2008). Mathematics intervention for first- and second-grade students with mathematics difficulties: The effects of Tier 2 intervention delivered as booster lessons. *Remedial and Special Education*, 29, 20–32. doi:10.1177/0741932507309712

Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N., Funk, C., Winter, A., . . . Pool, C. (2008). The effects of Tier 2 intervention on the mathematics performance of first-grade students who are at risk for mathematics difficulties. *Learning Disability Quarterly*, 31, 47–63. doi:10.2307/20528817

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18, 23–45. doi:10.1016/s0272-7358(97)00043-3

Dossey, J., McCrone, S. & Halvorsen, K. (2016). *Mathematics education in the United States 2016: A capsule summary book written for the thirteenth International Congress on Mathematical Education (ICME-13)*. Reston, VA: National Council of Teachers of Mathematics.

Dunn, T. J., Baguley, T., & Brunsden, V. (2013). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105, 399–412. doi:10.1111/bjop.12046

Fuchs, L. S., & Fuchs, D. (2006). A framework for building capacity for responsiveness to intervention. *School Psychology Review*, 35, 621–626.

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention: A decade later. *Journal of Learning Disabilities*, 45, 195–203. doi:10.1177/0022219412442150

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, 79, 1202–1242. doi:10.3102/0034654309334431

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

IBM Corp. (2015). *IBM SPSS statistics for windows* (Version 22.0). Armonk, NY: IBM Corp.

Jitendra, A. K., Dupuis, D. N., Star, J. R., & Rodriguez, M. C. (2016). The effects of schema-based instruction on the proportional thinking of students with mathematics difficulties with and without reading difficulties. *Journal of Learning Disabilities*, 49, 354–367.
doi:10.1177/0022219414554228

Jitendra, A. K., Harwell, M. R., Dupuis, D. N., & Karl, S. R. (2017). A randomized trial of the effects of schema-based instruction on proportional problem-solving for students with mathematics problem-solving difficulties. *Journal of Learning Disabilities*, 50, 322–336.
doi:10.1177/0022219416629646

Jitendra, A. K., Harwell, M. R., Dupuis, D. N., Karl, S. R., Lein, A. E., Simonson, G., & Slater, S. C. (2015). Effects of a research-based intervention to improve seventh-grade students' proportional problem solving: A cluster randomized trial. *Journal of Educational Psychology*, 107, 1019–1034. doi:10.1037/edu0000039

Jitendra, A. K., Harwell, M. R., Im, S-h., Karl, S. R., & Slater, S. C. (2018). Improving student learning of ratio, proportion, and percent: A replication study of schema-based instruction. *Manuscript submitted for publication*.

Jitendra, A. K., Harwell, M. R., Karl, S. R., Simonson, G. R., & Slater, S. C. (2017). Investigating a Tier 1 intervention focused on proportional reasoning: A follow-up study. *Exceptional Children*, 83, 340–358. doi:10.1177/0014402917691017

Jitendra, A. K., Star, J. R., Dupuis, D. N., & Rodriguez, M. C. (2013). Effectiveness of schema-based instruction for improving seventh-grade students' proportional reasoning: A

randomized experiment. *Journal of Research on Educational Effectiveness*, 6, 114–136. doi:10.1080/19345747.2012.725804

Jitendra, A. K., Star, J. R., Rodriguez, M. C., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction*, 21, 731–745. doi:10.1016/j.learninstruc.2011.04.002

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48, 281–355. doi:10.3386/w14723

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., . . . Busick, M. D. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms* (Publication No. NCSER 2013-3000). Washington, DC: National Center for Special Education Research. Retrieved from <https://ies.ed.gov/ncser/pubs/20133000/pdf/20133000.pdf>

Marshall, S. P. (1995). *Schemas in problem solving*. New York: Cambridge University Press.

Mayer, R. E. (1999). *The promise of educational psychology Vol. I: Learning in the content areas*. Upper Saddle River, NJ: Merrill Prentice Hall.

Meyer, J. P. (2011). jMetrik Version 2.1.0 [Computer software]. Retrieved from <http://itemanalysis.com>

National Center for Education Statistics (n.d.). *NAEP Data Explorer*. Washington, DC: Institute of Education Sciences, U.S. Department of Education. Accessed August 12, 2016. <http://nces.ed.gov/nationsreportcard/naepdata>

Özgün-Koca, S. A., & Altay, M. K. (2009). An investigation of proportional reasoning skills of middle school students. *Investigations in Mathematics Learning*, 2, 26–48. doi:10.1080/24727466.2009.11790289

Pearson (2004). *Group Mathematics Assessment and Diagnostic Evaluation [GMADE] (Level M, Form A)*. San Antonio, TX: Author.

Robinson, J. P. (2010). The effects of test translation on young English learners' mathematics performance. *Educational Researcher*, 39, 582–590. doi:10.3102/0013189x10389811

Ryoo, J. H., & Pullen, P. C. (2017). Regression discontinuity design for cross-sectional data, longitudinal data, and intervention research for special education. In J. M. Kauffman, D. P. Hallahan, & P. Cullen (Eds.), *Handbook of Special Education* (2nd ed., pp. 137–150). New York, NY: Taylor & Francis.

Schulte, A. (2016). Prevention and response to intervention: Past, present, and future. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (2nd ed., pp. 59–71). New York: Springer Science.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., . . .

Chen, M. (2012). Early predictors of high school mathematics achievement.

Psychological Science, 23, 691–697. doi:10.1177/0956797612440101

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Sciences*, 17, 13–19. doi:10.1016/j.tics.2012.11.004

Smith, M. S. & Stein, M. K. (2011). *5 practices for orchestrating productive mathematics discussions*. Reston, VA: National Council of Teachers of Mathematics.

Steiner, P., Cook, T. D., Shadish, W. R., & Clark, M. H. (2010). The importance of covariate selection in controlling for selection bias in observational studies. *Psychological Methods*, 15, 250–267. doi:10.1037/a0018719

U.S. Department of Education. (2015). *Digest of Education Statistics 2015*. Retrieved from

<https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2016014>

Vaughn, S., Klingner, J. K., Swanson, E. A., Boardman, A. G., Roberts, G., Mohammed, S. S., &

Stillman-Spisak, S. J. (2011). Efficacy of collaborative strategic reading with middle school students. *American Educational Research Journal*, 48, 938–964.

doi:10.3102/0002831211410305

Vaughn, S., & Swanson, E. A. (2015). Special education research advances knowledge in

education. *Exceptional Children*, 82, 11–24. doi:10.1177/0014402915598781

What Works Clearinghouse. (2014). *Procedures and standards handbook* (Version 3.0).

Washington, DC: Author.

Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A. K., . . . Ogbuehi,

P. (2012). *Improving mathematical problem solving in grades 4 through 8: A practice guide* (NCEE 2012-4055). Washington, DC: National Center for Education Evaluation

and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Retrieved from <http://ies.ed.gov/ncee/wwc/PracticeGuide.aspx?sid=16>

Table 1. *Summary of Student and Teacher Demographic Information*

	Treatment						
	SBI		Control		Total		
	n	%	n	%	n	%	
Student Information⁺							
Age	<i>M (SD)</i>	12.60	0.6	12.55	0.4	12.57	0.4
Sex	Female	268	50.5	283	48.0	551	49.2
	Male	254	47.8	300	50.9	554	49.5
Race	Asian	21	4.0	34	5.8	55	4.9
	Black	49	9.2	53	9.0	102	9.1
	Hispanic	116	21.8	196	33.3	312	27.9
	Multiracial	18	3.4	19	3.2	37	3.3
	White	318	59.9	281	47.7	599	53.5
ELL	Yes	46	8.7	69	11.7	115	10.3
	No	476	89.6	514	87.3	990	88.4
SpEd	Yes	51	9.6	59	10.0	110	9.8
	No	471	88.7	524	89.0	995	88.8
Missing (age, sex, Race, ELL, & SpEd)		9	1.7	6	1.0	15	1.3
FRL	Yes	134	25.2	159	27.0	293	26.2
	No	33	6.2	85	14.4	118	10.5
	Missing	364	68.6	345	58.6	709	63.3
Teacher Information							
Sex	Female	31	91.2	18	72.0	49	83.1
	Male	3	8.8	7	28.0	10	16.9
Location	Suburban	29	86.3	23	92.0	52	88.1
	Urban	3	8.8	0	0	3	5.1
	Rural	2	5.9	2	8.0	4	6.8
Math courses taken	<i>M (SD)</i>	7.53	4.8	11.12	10.3	9.05	7.7
Education courses taken	<i>M (SD)</i>	3.85	3.8	5	10.8	4.34	7.5
Years experience in math	<i>M (SD)</i>	10.76	7.0	9.84	10.5	10.4	8.6
PD hours in math	<i>M (SD)</i>	19.47	13.3	21.4	23.5	20.3	18.1

Note. + is defined by students whose GMADE pretest scores were 4-14; FRL = students eligible for free or reduced priced lunch; ELL = English language learner; PD = professional development; SpEd = students qualified for special education services.

Table 2. *Examples of Instructional Features in Control Textbooks that Incorporated SBI Components*

Key SBI components	Control Textbook Instruction
Identify the problem type and connect the problem to already solved problems	Two textbooks superficially addressed this component using questions (e.g., “Is the question looking for the part, the percent, or the whole?”) that did not emphasize the reasoning involved for students to demonstrate understanding of the problem situation. None of the textbooks provided opportunities for making connections (e.g., How is the problem to be solved similar to or different from a previously solved problem?).
Represent the problem using visual representations	All textbooks included a variety of visual representations (e.g., ratio tables, bar graphs, strip/bar diagrams).
Estimate the answer and use multiple solution strategies	None of the textbooks provided opportunities for estimating the answer prior to solving it. All textbooks used a variety of strategies to represent information in the problem (e.g., make a table, draw a diagram, act it), but the use of alternate solution strategies to solve a problem was only present in one textbook (e.g., compare unit rates and use the cross products property to determine whether the relationship between two ratios is proportional).
Apply problem-solving procedures	The majority of textbooks used worked examples and included question prompts (e.g., “What characteristic do you look for in the table in order to decide whether the relationship is proportional?” “What are you being asked to solve?”) to guide students in solving problems. One textbook included a general problem-solving procedure (understand, plan, solve, check) for modeling the problem; however, opportunities to apply this practice were limited to few examples.
Monitor and reflect on the problem-solving processes	In one of the textbooks, teacher questions prompted students to reflect on the problem-solving process (e.g., “How can percent help you understand situations involving money?” “How is compound interest different from simple interest?”).

Table 3. *Fidelity of Implementation and Quality of Instruction Data*

Variable	SBI ^a <i>M (SD)</i>	Control ^b <i>M (SD)</i>	<i>p</i>	<i>ES</i>
<i>Procedural Fidelity</i>				
Identifies problem type	2.25 (0.65)	1.76 (0.93)	.020	0.63
Problem similar/different	1.74 (0.63)	0.64 (0.86)	<.001	1.50
Represents key information	2.59 (0.40)	2.52 (0.65)	.646	0.13
Estimates solution	2.31 (0.90)	0.12 (0.60)	<.001	2.78
Uses multiple strategies	2.31 (0.54)	1.76 (0.93)	.012	0.75
Provides complete solution	2.72 (0.41)	2.64 (0.70)	.610	0.15
Checks solution	1.79 (0.54)	0.72 (0.98)	<.001	1.41
Total score	15.71 (2.42)	10.16 (2.70)	<.001	2.18
<i>Quality of Instruction</i>				
Sets lesson purpose	2.43 (0.46)	2.56 (0.87)	.490	-0.20
Provides lesson closure	0.88 (0.37)	1.12 (0.97)	.254	-0.35
Manages instructional time	2.60 (0.57)	2.32 (0.95)	.193	0.37
Minimizes math errors	2.84 (0.44)	2.56 (1.00)	.204	0.38
Total score	8.75 (1.06)	8.56 (2.16)	.688	0.12

Note. Total possible points on the procedural fidelity = 21; Total possible points on quality of instruction = 12. SBI = schema-based instruction; Effect size (ES) was calculated as the two conditions' mean difference divided by the pooled standard deviation (Hedges & Olkin, 1985).

Here $\alpha = \frac{.15}{13} = .0115$ following the Dunn-Bonferroni procedure. ^a*j* = 34. ^b*j* = 25.

Table 4. *Intercorrelations among Variables by Mathematics Difficulties Status and Treatment Group*

(a) Intercorrelations among variables by students whose PPS pretest scores were 4-14

	Students with MD (PPS pretest scores were 4-9)					Students without MD (PPS pretest scores were 10-14)				
	1	2	3	4	5	1	2	3	4	5
1. PPS Pretest	—	.258***	.153*	.148*	.113	—	.290***	.211**	.230***	.279***
2. PPS Posttest	.221***	—	.563***	.274***	.433***	.308***	—	.576***	.422***	.482***
3. PPS Delayed Posttest	.266***	.594***	—	.252***	.342***	.336***	.643***	—	.377***	.352***
4. GMADE pretest	.222***	.312***	.240***	—	.345***	.194**	.209***	.122*	—	.416***
5. GMADE posttest	.132*	.439***	.480***	.201***	—	.383***	.564***	.560***	.315***	—

(b) Intercorrelations among variables by students whose GMADE pretest scores were 5-15

	Students with MD (GMADE pretest scores were 5-10)					Students without MD (GMADE pretest scores were 11-15)				
	1	2	3	4	5	1	2	3	4	5
1. PPS Pretest	—	.547***	.478***	.130	.359***	—	.530***	.476***	.144*	.388***
2. PPS Posttest	.559***	—	.645***	.133	.513***	.532***	—	.609***	.162*	.459***
3. PPS Delayed Posttest	.601***	.707***	—	.074	.460***	.586***	.732***	—	.205**	.383***
4. GMADE pretest	.080	.135*	.018	—	.038	.269***	.211***	.131*	—	.195*
5. GMADE posttest	.409***	.564***	.562***	.046	—	.463***	.564***	.590***	.191***	—

Note. PPS = proportional problem solving; GMADE = Group Mathematics Assessment and Diagnostic Evaluation; MD = mathematics difficulties; Correlations for the control group are above the diagonal; those for the SBI group are below the diagonal; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 5. *Descriptive Statistics for Outcome Variables by Treatment*

Variable	by PPS pretest scores (35 th percentile cut-off = 9 or 10)						
	SBI			Control			
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>ES</i>
PPS posttest	129	14.02	4.77	125	11.22	4.02	0.64**
PPS delayed posttest	124	12.86	4.64	121	10.93	4.39	0.43**
GMADE posttest	128	12.32	4.11	119	11.56	3.77	0.19
by GMADE pretest scores (35 th percentile cut-off = 10 or 11)							
	SBI			Control			
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>ES</i>
PPS posttest	138	14.21	5.57	104	11.49	4.55	0.53**
PPS delayed posttest	129	12.94	5.21	94	10.93	4.80	0.40**
GMADE posttest	137	12.91	4.18	100	11.47	3.73	0.36**

Note. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 6. *Regression Discontinuity for the PPS and GMADE Outcome Variables*

Outcomes	by Bandwidth of PPS pretest scores					by Bandwidth of GMADE pretest scores				
	± 1 (8 ~ 10)	± 2 (7 ~ 11)	± 3 (6 ~ 12)	± 4 (5 ~ 13)	± 5 (4 ~ 14)	± 1 (9 ~ 11)	± 2 (8 ~ 12)	± 3 (7 ~ 13)	± 4 (6 ~ 14)	± 5 (5 ~ 15)
PPS posttest										
β	2.11	2.405	2.337	2.395	2.392					
Stand. β	0.225	0.246	0.238	0.240	0.233					
p	< .001	< .001	< .001	< .001	< .001					
n	376	595	760	905	1,024					
PPS delayed posttest										
β	1.165	1.054	0.986	1.099	1.106					
Stand. β	0.131	0.117	0.107	0.116	0.113					
p	0.012	0.005	0.004	0.001	< .001					
n	363	570	726	862	973					
GMADE posttest										
β						1.497	1.165	1.009	1.128	1.132
Stand. β						0.178	0.138	0.118	0.130	0.129
p						0.001	0.001	0.001	< .001	< .001
n						339	573	735	903	1,039

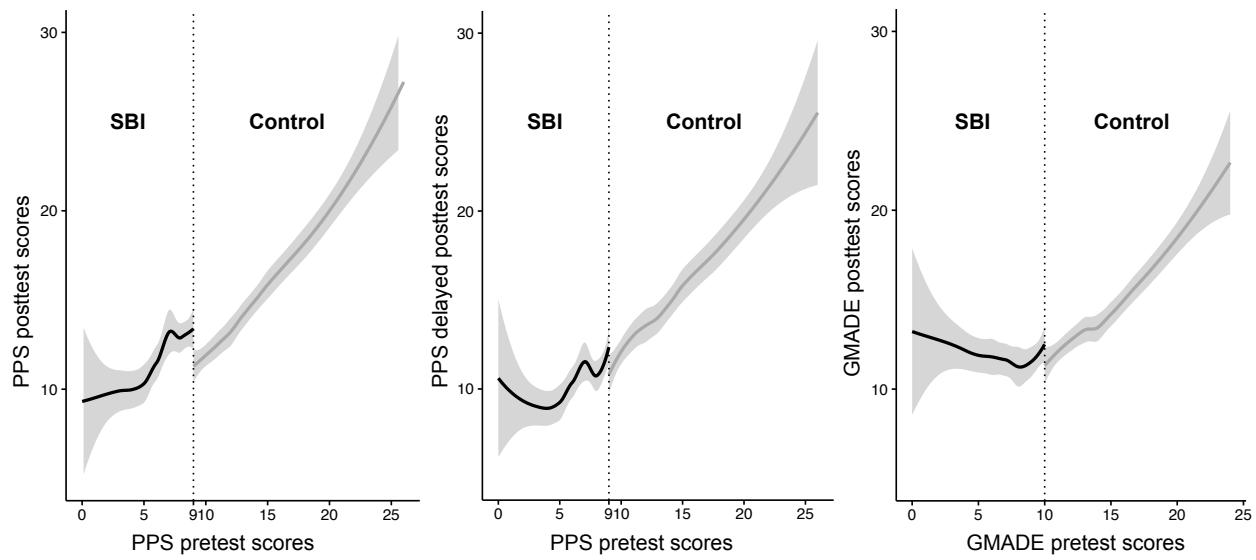


Figure 1. The effect of SBI on PPS posttest (left), delayed posttest (middle), and GMADE posttest scores (right) by PPS or GMADE pretest scores. The vertical distance between the solid lines as they approach the threshold (i.e., vertical “jump” in outcome variables) is the regression-discontinuity-based effect estimate. The gray shading represents the 95% confidence interval around the line of best fit.

Appendix A

Schema-Based Instruction (SBI): Sample Lessons (abbreviated)

Solving Ratio Word Problems (Lesson 3)

Teacher: Today, you will learn to apply ratios in a variety of contexts to solve word problems. Remember, you learned that a ratio expresses a multiplicative relationship between 2 quantities in a single situation. Also, there are many types of ratios (such as part-to-whole, part-to-part, and whole-to-part) that can describe a single situation.

Problem: On Thursday, the cafeteria at Osseo Middle School sold: 42 smoothies, 75 main line lunches, 80 cookies, 51 bags of chips, 100 salad bar lunches, and 26 breakfast bars. What is the ratio of the number of main line lunches sold to the number of salad bar lunches sold on Thursday?

Teacher: There is a lot of information given in this problem. Look at all those numbers! What must you solve for in this problem?

Students: *The ratio of the number of main line lunches sold to the number of salad bar lunches sold.*

Teacher: Excellent! Do we need information about the number of smoothies or the number of cookies or any of the other things that were sold? (No) So, what do we know that will help us answer this question?

Students: *We know that they sold 75 main line lunches and 100 salad bar lunches.*

Teacher: Yes! What about the 51 bags of chips?

Students: *We don't need that information to solve this problem.*

Teacher: Well done. Sometimes, there is extra information in problems that we don't need to solve the problem. So, what is the ratio of the number of main line lunches sold to the number of salad bar lunches sold?

Students: *75:100.*

Teacher: Excellent! Why wouldn't the correct answer be 100:75?

Students: *Because the question asks about the ratio of the number of main line lunches to the number of salad bar lunches.*

Teacher: Great! Note that the ratio 100:75 *does* tell us something about this situation; it tells us the ratio of the number of salad bar lunches to the number of main line lunches. However, this particular ratio is not what the problem asked us. The correct answer here is 75:100. How do we find the ratio value?

Students: *We divide the front term by the back term. It is $75 \div 100 = 0.75$.*

Teacher: Remember, a ratio is a multiplicative relationship between two quantities in a single situation (such as Thursday at the Osseo Middle School cafeteria). There can be many ratios that correctly describe a single situation. Let's look at the information given in this problem again. What other ratios can you think of that correctly describe this situation?

Students: *The number of smoothies to the number of cookies (80:42); the number of cookies to the number of bags of chips (80:51); the number of bags of chips to the number of breakfast bars (51:26); the number of breakfast bars to the number of main line lunches (26:75); the number of smoothies to the number of bags of chips (42:51); the number of bags of chips to the number of smoothies (51:42).*

Teacher: Excellent! (If students don't come up with a part-to-whole ratio, then ask, "What would be an example of a part-to-whole ratio in this problem?")

Students: *You could compare the number of smoothies sold to the total number of food options sold (42:374).*

Solving Proportion Word Problems (Lesson 6, Problem 2)

Teacher: Today we will learn to apply rates/ratios in a variety of contexts to solve proportion word problems. Remember, you learned that a rate is a comparison of any two quantities with different units (e.g., passengers to a minivan) that expresses a multiplicative relationship. Rates tell how many of one thing there are for a certain number of another thing (35 passengers to 5 minivans). Next, you will learn to solve proportion problems that involve two pairs of rates/ratios that are equivalent. A *proportion* is a statement of equality between two ratios/rates that allows us to think about the ways that two situations are the same. Often, there is an *If-Then* statement that tells about two ratios/rates that are equivalent.

Here are 4 steps we will use to solve proportion word problems:

D	Discover the problem type
I	Identify information in the problem to represent in a diagram(s)
S	Solve the problem
C	Check the solution

Let's use these 4 steps (DISC) to solve this problem.

**Tuna Casserole Recipe
(16 servings)**
2 cups cheese
1 cup milk
4 cups peas
2 cans tuna
1 cup noodles

Problem:

If Dora needs to make the casserole for 40 people at her party next weekend, how many cans of tuna does she need?

Teacher: To discover the problem type, I will read, retell, and examine information in the problem to recognize the problem type. The information for this problem is given in the recipe table and in the text. To understand this problem, first read the information given in the text and then

the table when you retell this problem. What is the information given in the text and what are you asked to solve?

Students: I know that Dora will serve casserole to 40 people at her party next weekend. I don't know how many cans of tuna she will need to serve these 40 people – this is what I need to solve in the problem.

Teacher: Good! The recipe table gives you information about the ingredients needed to make a casserole for 16 people. From this table, what information do you need to solve for the number of cans of tuna needed to serve 40 people? Explain.

Students: The number of people and cans of tuna, because the problem is comparing number of people (or servings) to cans of tuna.

Teacher: Excellent. From the recipe table, we know that 2 cans of tuna are needed to make a casserole for 16 people. What kind of a problem is this? How do you know?

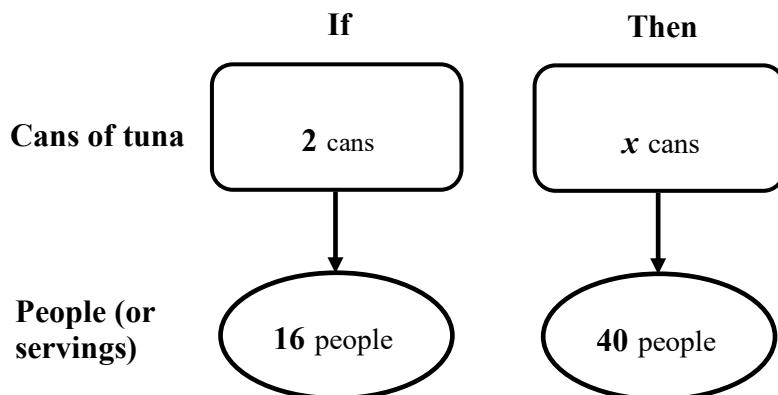
Students: It is a proportion problem, because it is a statement of equality between two ratios/rates that allows us to think about the ways that two situations are the same. The If statement in this problem describes the first situation - a ratio between two quantities (2 cans of tuna to 16 people or servings) and the Then statement describes the second situation - with a multiplicative increase in the quantities (x cans of tuna to 40 people) so that the two ratios will be the same when we solve for x.

Teacher: Excellent! The ratio in the first situation ($\frac{2 \text{ cans of tuna}}{16 \text{ people}}$) can be applied to the second situation ($\frac{x \text{ cans of tuna}}{40 \text{ people}}$). That is, if 2 cans of tuna serve 16 people, then how many cans of tuna are needed to serve 40 people? Is this problem different from/similar to another problem you have already solved?

Students: This problem is different from the first proportion problem, because the information is presented in a table. It is similar to the first problem, because it describes an If-Then statement of equality between two ratios, with two ratios comparing two quantities (cans of tuna-to-people) in two separate situations.

Teacher: You figured this is a proportion problem. Now identify information in the problem to represent in our proportion diagram.

....



....

Now you are ready to solve the problem. Before you solve the problem, first come up with an estimate for the answer. What is the estimate of the number of cans of tuna Dora needs to serve 40 people if 2 cans of tuna serve 16 people? Explain.

Students: It is hard to come up with an estimate, but I do know a few things about the answer. I know that the answer is more than 2 cans, because Dora needs 2 cans to serve 16 people, and she needs to serve more than 16 people. It is hard to know how much more, but I might guess that she needs to double the recipe. So a little more than 4 cans of tuna is a good estimate here, because 4 cans would serve 32 people and we need to serve 40 people.

Teacher: Excellent estimate! What is the math equation when you translate the information in the diagram?

Students: $\frac{2 \text{ cans of tuna}}{16 \text{ people}} = \frac{x \text{ cans of tuna}}{40 \text{ people}}$

Teacher: What strategy will you use to solve for the number of cans of tuna (which is x) it would take to serve 40 people.

Students: The unit rate strategy.

Teacher: Show the steps you use to solve the problem using the unit rate strategy and give me the answer.

Students: Because we need to find the number of cans of tuna (the x is in the numerator) to serve 40 people, we have to reason up to analyze the relationship between the two quantities as shown below:

$$\frac{2 \text{ cans of tuna}}{16 \text{ people}} = \frac{x \text{ cans of tuna}}{40 \text{ people}}$$

In this situation, I start with the given ratio and look to see what number I can multiply the numerator by to get the denominator. So, 2 times what number equals 16? It is 8, since $2 \cdot 8 = 16$. The 8 tell us the unit rate: 1 can of tuna serves 8 people. So, to solve for x , we apply this unit rate to the other ratio. When we multiply x cans of tuna by 8, we should get 40. What number times 8 equals 40? It is 5, since $8 \cdot 5 = 40$. So, 5 cans of tuna will serve 8 people.

Teacher: What do you do next?

Students: Check if the answer makes sense.

Teacher: How do you check the solution?

Students: Look back to see if our estimate is close to the exact answer.

Teacher: You estimated the answer to be more than 4 cans. The answer to this problem is 5 cans of tuna. So, I think your estimate (more than 4) is a good prediction. Now check to see if the answer makes sense. Does 5 seem right? Explain.

Students: Yes, if 2 cans of tuna serve 16 people, then 5 cans of tuna serve 40 people. This seems right, because 5 cans should serve more people than 2 cans of tuna. So, the answer 5 cans of tuna, which is more than 2 cans of tuna seems right. I will also check the answer by taking the ratio in the Then situation

(5 cans of tuna) and seeing if the value of this ratio is equal to the ratio in the If situation $\frac{5 \text{ cans of tuna}}{40 \text{ people}}$

(2 cans of tuna). When I simplify both ratios, I get $\frac{1}{8}$ can of tuna per person (or 8 people per $\frac{16}{8}$ people can of tuna--both of these are the unit rate), which tells me that they are equivalent.

Teacher: Great! In this problem, Dora had a specific reason for using 5 cans of tuna for 40 people. She wanted to make sure that the tuna casserole she made for 40 people would taste the same (e.g., would have the same amount of tuna) as the casserole recipe for 16 people. Dora wanted the ratio of the number of cans of tuna to the number of people to be the same for both situations (2 cans of tuna and 5 cans of tuna.) This is a proportion problem, because it is a statement of equality between two ratios that allows us to think about ways that two situations are the same.

Sample Challenge Problems from the SBI program

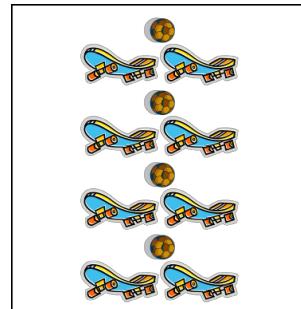
1. There are 28 students in Kendra's physical education class. On the fitness test, one fourth of the students did 50 pushups, one half of the students did 25 pushups, one-fourth of the students did 13 pushups. What was the total number of pushups that Kendra's class did? What is the average number of pushups a student in Kendra's class did?
2. You are playing with a deck of playing cards. You place 6 cards on the table *face up* and 1 card on the table *face down*, for a total of 7 cards. You want a ratio of 1 card *face down* to 4 total number of cards (1:4). How many *face down* cards do you need to add?
3. Sonja is making chili for a contest. She needs to make four times her regular recipe. Her recipe calls for 2 cups of diced tomatoes. She has 5, 16-ounce cans of tomatoes. Does Sonja have enough cans of tomatoes to make the quadrupled chili recipe? How many cups and cans of tomatoes does she need? (*Note: 1 cup = 8 ounces*)
4. Brian works at a sporting goods store that marks up its merchandise 40% from what it cost the store. Brian can buy things for 10% below cost (what it cost the store). He's interested in a skateboard that shoppers at the store would have to pay \$60 for. As a store employee, how much would Brian need to pay for the skateboard?

Appendix B**Sample Items from the Proportional Problem-Solving Test**

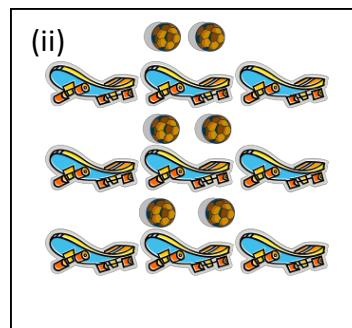
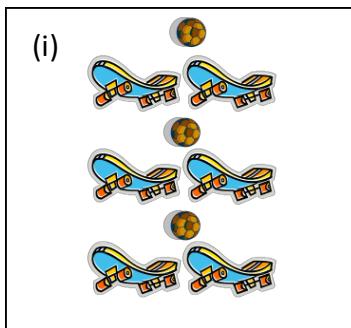
1. In making a garden fertilizer, a gardener mixes 2 kg of a nitrate, 3 kg of a phosphate, and 6 kg of potash. What is the ratio of nitrate to the total amount of fertilizer?
 - A. $\frac{11}{9}$
 - B. $\frac{2}{3}$
 - C. $\frac{2}{9}$
 - D. $\frac{2}{11}$
2. Which of the following ratios is equivalent to the ratio of 6 to 4?
 - A. 12 to 8
 - B. 4 to 6
 - C. 2 to 3
3. At the school carnival, Carmen sold 3 times as many hot dogs as Shawn. The two of them together sold 152 hot dogs. How many hot dogs did Carmen sell?
 - A. 38
 - B. 51
 - C. 114

4. Skateboard-Soccer Ball Problem

The picture to the right shows several skateboards and soccer balls.



There are skateboards and soccer balls shown below in boxes (i) and (ii). The ratio of skateboards to soccer balls in box (i) **is equivalent** to the ratio of skateboards to soccer balls in the box above. The ratio of skateboards to soccer balls in (ii) **is not equivalent** to the ratio of skateboards to soccer balls in the box above. Explain why the ratio in box (ii) **is not equivalent** to the ratio in the box above.



5. If 2 cups of flour are needed to make one dozen cookies, how many cups of flour will be needed to make 18 cookies?

- A. 3
- B. 36
- C. 4
- D. 6

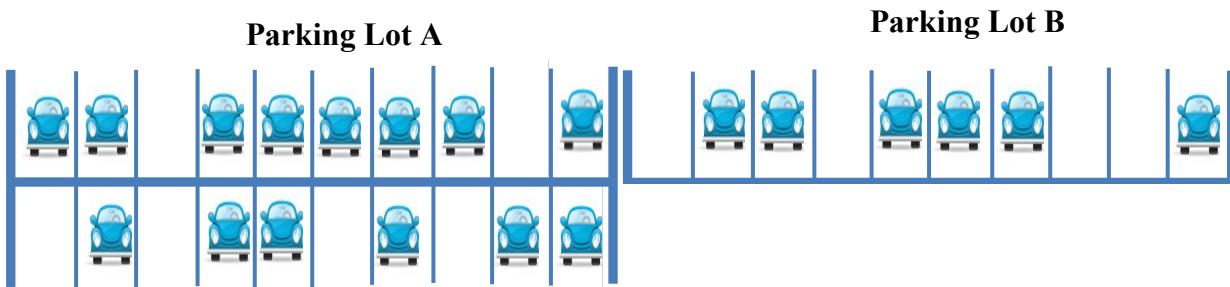
6. Some classmates compared their scores on a recent math test.

- Molly answered 15 out of every 20 questions correctly.
- Brittany answered 7 out of every 8 questions correctly.
- Desiree answered 7 out of every 10 questions correctly.
- Nick answered 4 out of every 5 questions correctly

Which student answered more than 80% of the questions correctly?

- A. Molly
- B. Brittany
- C. Desiree
- D. Nick

7. Below are pictures of two parking lots at the Mall of America, showing which parking spaces are full and which are empty.



John says that parking lot A is emptier than parking lot B, because parking lot A has 6 empty spaces while parking lot B has only 4 empty spaces.

Do you agree or disagree with John?

Explain your answer in 1–2 sentences.

8. In the model town that a class is building, a car that is 15 feet long is represented by a scale model 3 inches long. If the same scale is used, a house 35 feet high would be represented by a scale model how many inches high?

A. $\frac{45}{35}$
B. 5
C. 7
D. $\frac{35}{3}$

9. Lisa and Tracy have both gotten taller this year. Last year, Lisa was 45 inches tall, but now she is 51 inches tall. Last year Tracy was 40 inches tall, but now she is 46 inches tall. Who had the greatest percent of increase in their heights, Lisa or Tracy?

a. Tracy has the greatest percent of increase in her height.
b. Lisa has the greatest percent of increase in her height.
c. Both girls have the same percent of increase in their heights.